Computational Neuroscience
Dr. Sharba Bandyopadhyay
Department of Electronics and Electrical Communication Engineering
Indian Institute of Technology Kharagpur
Week – 05
Lecture – 21

Lecture 21 : Random variables and random process

Welcome. We have been discussing about the process of spiking. We have discussed how to generate spikes or rather how to model the spiking behavior in single neuron models. We have understood the process of spiking. We have understood various nuances related to spiking behavior in single neuron models. So now it is time to move on to the next phase of our course and which has to do with understanding how a series of action potentials or spikes encode information about the stimulus or about the external world or about a certain motor action that is being performed or about a certain thought and so on or a variety of cognitive aspects are represented in spike trains.

So the very first idea that we need to understand, need to remember throughout the course is that the neuronal responses that we will talk about or the spike trains that we will talk about are in general random in nature. That is not totally random but it is not repeatable from trial to trial. In other words, there is an associated stochasticity involved in neuronal spiking. So if you remember, recollect, let us try to see that we have a dendritic branch.

So if we are recording from this particular neuron, let us say extracellularly with, by getting spikes over time in this manner, let us say, so we can represent it as events over time that are occurring while some other thing, associated thing may be going on in parallel and we try to relate this particular spike train that we are observing with whatever is going on in parallel in the external world or even internal processes. So if we, these spike trains, if you remember, is generated based on what is going on intracellularly based on current injections through synapses or inputs onto the dendrites or even on the soma. So there can be thousands of synapses. So internally there may be a number of things going on only when there are threshold crossings that we get to see spikes like this. So if this is the membrane potential V_M and let us say this is V_T, then only here do we get the spikes.

So the first source of this variability, that is, let us say the event, whatever is going on in parallel is repeated, we may or may not get a spike at this particular time or may or may not get a spike at all over this interval or may get the spike more than one spike within the same interval. So there are a number of possibilities. Now what makes this variability occur? It is primarily, it is the unreliability in the synapses or synaptic transmission. We already know about the stochasticity of ion channels in the level of ion channels that may or may not allow the spike to happen at that particular time point. Then there is, there is hierarchically these same variabilities gradually add up over the different layers of neurons that we are going through and finally what we get is something that is akin to what we tend to call a Poisson process.

So we will get into the details of these in a little bit but what we want to first know that in order to study these spike trends, we come up with a way to quantify these spike trends or treat them mathematically and that treatment can occur at variety of scales of time. That is I may consider the activity over very small time windows. Let us say this is the spike trend. We may consider the activity over very small time windows over time like this being delta. We may consider the activity over medium sized time windows.

This is another delta and we may consider the activity over a very large time window. In that case that whole thing is delta. So our measure of the response or quantification of the spike trend depends on how we discretize or how we treat the time interval in general when we are talking of coding or decoding by neurons. So in order to look into these aspects, the first thing that we need to recollect or have a little primer about is the treatment of random variables and we will then go on to discuss what we will call random processes. So if we ask the question what is a random variable, we will denote it by this.

The random variable is essentially a mapping from a sample space or a set of events to the real number line or some other metric space. For us let us consider that it is a real number line and there is a probability associated with each of those values that are taken by the random variable. So let us say this is the set of possible outcomes of a random experiment. So from there if we consider the real number line 0 up to plus infinity minus infinity here, then this mapping on to the real number line of the outcomes of the random experiment, this mapping is essentially what is a random variable. And so since the outcomes of the random experiment, each of them have a certain probability associated with them, the values taken by the random variable also have the same probability associated with them.

So for example we can have a random variable X let us say which takes on the value 0 and 1 when we say that it is a coin tossing experiment where we denote 0 as the outcome head and 1 as the outcome tail, then this X is a random variable taking on value 0 and 1 has probabilities of taking on the value probability of

$X = 0$ is probability of the occurrence of head in the coin toss experiment which can be P or if it is an unbiased coin it is half which can be half when unbiased. And similarly we have the probability of $X = 1$ equals probability of tail equals 1 minus P or in the case of unbiased coin it is half. So this is just a recollection of what a random variable is and along with it we remember that the random variable X can be discrete as above that is it takes on a discrete set of values like in the case of head and tail or if we consider the random experiment as a die throw where the value of the random variable taken is the number of dots that appear on the top of the die after it settles down then it is a discrete random variable taking on 6 possible values and each of them have certain probabilities in the unbiased die case they are 1/6. So you are aware of these things. So in the case of discrete random variables what defines the random variable is its probability mass function or probability distribution so PMF and we are given that probability of X equal to X_i is let us say P_i and your i varies from 1 up to capital N so there are capital N possible values taken by X and we also know that summation of P_i i equals 1 to capital N need to sum to 1 this is from basic probabilities that we know.

So whatever that there is to be known about the random variable X is given in the PMF. So from this we can determine what is the average value of the random variable. So if we have the given values X_i of 1 to N i equals 1 to N we say the expected value or the average value of the random variable X is simply summation of P_i times X_i i equals 1 to capital N and similarly we extend the idea to the variance where we talk about the expectation of X squared minus expectation of X squared I am sorry expectation of X minus expectation of X whole squared this is the variance of the random variable X variance of the random variable X. So once we have the PMF so this is the expectation of X and so once we have the PMF we can actually get the expected value or average value of variety of different functions of the random variable X of which one of them is variance. So based on these ideas of random variable is some is what is built is stochastic processes or random processes where a random process is essentially a sequence of random variables.

So we also often tend to say that it is random variable at each and every time instant in case of continuous time random processes. So let us think of a way in which we are measuring some phenomena and somehow we could see this phenomena multiple number of times in a thought experiment that we are observing the same thing once again at the same time but once again and yet again and so on. So what we mean by that it is let us say there is a time 0 or relative to which we are measuring time then and the random process is taking on or the observation that we are having have different values at the different time points let us say this is T_i. So we are observing this yet again and again we come to the time point T_i. Now what we are saying that at the new T_i the value taken by the random process or X T_i the random variable X T_i is simply another outcome of the experiment that goes behind creation of the random variable X at time point T_i.

So these different observations of the same phenomena multiple number of times are essentially called realizations of a random process. Now that means that at this time point T_i we are essentially observing the random variable X T_i multiple number of times that is we are doing the same experiment again and again and we are getting a certain value associated with a certain probability that is occurring each and every time or in each and every realization. So if we were to be able to observe this random process in finite number of times in finite realizations then we could characterize the random variable X T_i completely and get its distribution. In this case let us say if the X T takes on values that are continuous over an interval on the real number line then this X T is a continuous random variable for this T equals particular T_i. And so when we have a continuous random variable instead of a PMF we describe it based on a probability density function or PDF such that it is given by if let us say probability of X that it takes on the value X_i is we cannot actually write it for a continuous random variable this will be 0.

So we always consider probability of a continuous random variable X to take on a value between let us say greater than X_i and X is less than equal to X_i plus ΔX or we can drop the suffix here let us write this as X and less than equal to X plus ΔX. So this probability can be written as integral of P_X into P_X dX which we are integrating from X to X plus ΔX where we can define let us say it is some particular X_0 to X 0 plus ΔX. So in this case we since X can take on values over the continuous interval we define I mean the random variable is defined by the PDF and if we have a random process which takes on continuous values then it is also called a continuous random process then each of those time points T_1, T_2 are essentially for each of those T we have a particular PDF or density function of X_T. So the probability density function of X_T needs to be known in order to know about the value taken by the random variable at time T. So this statement is not entirely complete because now that we have a random process over time there are multiple random variables that we have that are preceding or occurring after the time point T_i.

Let us say this is X T_j and then it is let us say that T_k it is X T_k. So the value at time point T_i may be dependent on previous values and in a theoretical way can depend on the past values. In reality it cannot I mean in future values in reality it cannot really depend on future values but if we think of the independent variable time as something else we can theoretically think that it is dependent on values to the right as well as to the left. So this brings us the idea of the dependence of the distribution or the probability density function of X_T and it depends on the multiple values taken by the random variable or the random process at the different time points in that realization. So in a sense we require a joint distribution of X_T over all the time instances together.

So we will leave it up to here in terms of the ideas of development of the random process and we will continue on to describe a few other things that have to do with the measures of these random processes and that is the functions that we need or the items that we need in future are what we call the autocorrelation function and the cross correlation function. So as we were saying we have this random process X_T the autocorrelation function is simply defined as the expected value of the random variable X_{T_1} and X_{T_2} product of these two and that is called $R_{X_{T_1},X_{T_2}}$ or we can also write it as R_X at T_1, T_2 and that is this expected value. Now we will see that I mean under certain conditions which we call wide sense stationarity or strong stationarity in both cases this $R_{X_{T_1},X_{T_2}}$, $R_{X_{T_1},X_{T_2}}$ can be changed to $R_X(\tau)$ of a single variable which is the difference in time between T_1 and T_2. So we had this random process over time let us say this is our T_1, this is our T_2 and let us say this difference in time is τ. So under the condition of stationarity or wide sense stationarity or weak stationarity we can say that this expectation is independent of the location of T_1 and T_2 but is only dependent on the difference between T_1 and T_2.

So the expected value of the process at T_1 and T_2 is the same as if we translate T_1 by some capital T and T_2 by some capital T that is T_2 plus capital T and this point is T_1 plus capital T then the expected value of these two points are the same as these two and so on. So essentially what is meant by stationarity is that the functions of the random process of at different time points are independent of translations in time. So that means that our expected value can be obtained from by observing only those two time points only what will matter is the difference in time points of the two random values that we are taking. Now in order to compute this expectation X_{T_1}, X_{T_2} or in the case of stationarity expectation of now we can write it as X_T and $X_{T-\tau}$ where now we are representing any time point T because it is independent of the position T_1 and T_2 it only depends on the time the gap τ and this requires knowing the as we know knowing the joint distribution at X_T and $X_{T-\tau}$ or rather at time point T and time point $T-\tau$. There is a further assumption that can be used which are suitable for our cases also or that is we have a large amount of data and that is if we want to find out what this expectation is we can do that by observing the random process over long or rather infinite period of time and this expectation can be computed by doing a time average instead of doing an average of X_T and $X_{T-\tau}$ product over the probability density.

In other words what we mean is that this can be computed by simply integrating this X_T and $X_{T-\tau}$ over time from whatever is the range of values minus infinity to infinity and that is what we will generally call $R_{XX}(\tau)$ or the auto correlation function of the random variable X. Similarly we have cross correlation function where there are two random processes X and Y that are being observed and the time point T_2 is being taken on another random variable and T_1 on one of the random variables and again under the conditions of joint stationarity of X and Y and our gordicity we can then write that the cross correlation function $R_{XY}(\tau)$ is nothing but the integral of minus infinity to infinity $X_T Y_{T-\tau} dt$ and this can be used to actually perform a lot of calculations to do with neuronal processing in the conversion from stimulus to the response. So to add on this that the random process and random variables so we saw that the random variables we treat can have discrete and later on we saw they may be also continuous in these two cases and they are defined by the PMF on this case this is defined by the PDF. Similarly in the random processes case the values taken just like the random variables can be discrete and also can be continuous we described everything with as if things are continuous but they can be discrete as well and the other aspect is these each of these may be combined with what is happening on time that is discrete time or continuous time discrete time and continuous time. So in other words the random process may be taking on discrete possible values and at discrete time intervals discrete time points can take on discrete possible values continuously on time x that is a continuous time discrete valued random process and it can be continuous random process in discrete time or continuous valued random process in continuous time all of these are possible and some of these will come up in our case.

So we have said that the response I mean we introduced the idea of the random variable and random processes by talking about the variability in responses. Similarly we can think of the stimulus or the external world or whatever is driving the neuron as also a random process or if it takes on or random variable depending on the situation and this relationship between these two random events or random variables or random processes as the case may be is what we will be actually handling in our discussions of encoding and decoding by neurons. So in the next lecture we will start from here with these ideas as a background and there will be more material regarding them in your handouts in your reading material in more details and so we will go on to understand what or how we understand the relationship between the stimulus and the response. So in general we will call the other side stimulus and the response space of course is the particular random variable that we had been discussing. So we will end here and start off with what we mean by response space, response measure and also what is the nature of the statistics of the response measure or spike trends in the next lecture. Thank you.